A Formal Characterization of the Outcomes of Rule-Based Argumentation Systems

Leila Amgoud and Philippe Besnard

IRIT – CNRS
118, route de Narbonne
31062, Toulouse Cedex 09

Abstract. Rule-based argumentation systems are developed for reasoning about defeasible information. As a major feature, their logical language distinguishes between strict rules and defeasible ones. This paper presents the first study on the outcomes of such systems under various semantics such as naive, stable, preferred, ideal and grounded. For each of these semantics, it characterizes both the extensions and the set of plausible inferences drawn by these systems under a few intuitive postulates.

1 Introduction

There are two major categories of instantiations of Dung's abstract argumentation framework [4]. A category uses *deductive logics* (such as propositional logic [2,6] or Tarskian logics [1]). The second category uses *rule-based languages* [3,5,7] which distinguish between *facts*, *strict rules* (they encode strict information), and *defeasible rules* (they describe general behavior with exceptional cases). Despite the popularity of rule-based argumentation systems, the results they return have not been characterized yet. The following questions are still open:

- what are the underpinnings of the extensions under various semantics?
- do Dung's semantics return different results as at the abstract level?
- what is the number of extensions a system may have under a given semantics?
- what are the plausible conclusions with such systems?

In this paper, we answer all the above questions. We start with a knowledge base called a *theory* (a set of facts, a set of strict rules and a set of defeasible rules), we define a notion of a *derivation schema* which we use to generate arguments from the theory. For the sake of generality, the attack relation is left unspecified. We extend the list of postulates proposed in [3] with three new postulates. We investigate outputs of rule-based argumentation systems that satisfy all the postulates. We show that naive extensions return maximal *options* of the theory (an option being a sub-theory that gathers all the facts and strict rules, and a maximal -up to consistency- set of defeasible rules that do not conflict with the strict part). Every maximal option gives birth to a naive extension. Furthermore, the set of plausible conclusions under the naive semantics contains all the conclusions that are drawn from all the maximal options. Stable extensions return maximal options but not necessarily all of them, it depends on the attack relation at work. Should not all maximal options be picked as stable extensions, defining an attack relation that discard exactly the spurious ones turns out be tricky. The same results hold for preferred semantics. We characterize both ideal and grounded extensions.

2 Rule-Based Argumentation Systems

In what follows, we consider the language used in [3]. Let $\mathcal{L}$ is a set of *literals*, i.e., atoms or negation of atoms. The negation of an atom x from $\mathcal{L}$ is denoted $\neg x$. Three kinds of information ($x, x_1 ... x_n$ denoting literals in $\mathcal{L}$) are distinguished:

- *Facts*, which are elements of $\mathcal{L}$
- *Strict rules*, which are of the form $x_1, \ldots, x_n \rightarrow x$
- *Defeasible rules*, which are of the form $x_1, \ldots, x_n \Rightarrow x$

Throughout the text, rules are named $r_1, r_2, \ldots$ For each rule $r = x_1, \ldots, x_n \rightarrow x$ (as well as $r = x_1, \ldots, x_n \Rightarrow x$), the *head* of the rule is $\mathtt{Head}(r) = x$ and the *body* of the rule is $\mathtt{Body}(r) = \{x_1, \ldots, x_n\}$. A strict rule expresses general information that has no exception, e.g. “penguins cannot fly” whereas a defeasible rule expresses general information that may have exceptions, e.g. “birds can fly”.

Definition 1 (Theory). *A* theory *is a triple $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$ where $\mathcal{F}$ is a set of facts and $\mathcal{S}$ (resp. $\mathcal{D}$) is a set of strict (resp. defeasible) rules.*

Notation. Let $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$ and let $\mathcal{T}' = (\mathcal{F}', \mathcal{S}', \mathcal{D}')$ be two theories. We say that $\mathcal{T}$ is a *sub-theory* of $\mathcal{T}'$, written $\mathcal{T} \sqsubseteq \mathcal{T}'$, iff $\mathcal{F} \subseteq \mathcal{F}'$ and $\mathcal{S} \subseteq \mathcal{S}'$ and $\mathcal{D} \subseteq \mathcal{D}'$. The relation $\sqsubset$ is the strict version of $\sqsubseteq$ (i.e., it is the case that at least one of the three inclusions is strict).

The notion of consistency is defined as follows.

Definition 2 (Consistency). *A set $X \subseteq \mathcal{L}$ is* consistent *iff $\nexists x, y \in X$ s.t. $x = \neg y$. It is* inconsistent *otherwise.*

Assumption 1. *The body of every (strict/defeasible) rule is* finite *and* not empty. *Moreover, for each rule r,* $\mathtt{Body}(r) \cup \{\mathtt{Head}(r)\}$ *is consistent. We say that r is consistent.*

The notion of a *derivation schema* generalizes derivations as defined in [5,8] and others. It shows how literals can follow from a theory.

Definition 3 (Derivation schema). *Let $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$ be a theory and $x \in \mathcal{L}$. A* derivation schema *for x from $\mathcal{T}$ is a finite sequence $d = \langle (x_1, r_1), \ldots, (x_n, r_n) \rangle$ s.t.*

- $x_n = x$
- *for $i = 1 \ldots n$,*
  - $x_i \in \mathcal{F}$ *and* $r_i = \emptyset$, *or*
  - $r_i \in \mathcal{S} \cup \mathcal{D}$ *and* $\mathtt{Head}(r_i) = x_i$ *and* $\mathtt{Body}(r_i) \subseteq \{x_1, .., x_{i-1}\}$

$\mathtt{Seq}(d) = \{x_1, \ldots, x_n\}$.
$\mathtt{Facts}(d) = \{x_i \mid i \in \{1, \ldots, n\}, r_i = \emptyset\}$.
$\mathtt{Strict}(d) = \{r_i \mid i \in \{1, \ldots, n\}, r_i \in \mathcal{S}\}$.
$\mathtt{Def}(d) = \{r_i \mid i \in \{1, \ldots, n\}, r_i \in \mathcal{D}\}$.

Notation. In order to improve readability, we somehow abuse the notation in derivation schemata: We use the name of the rules instead of the rules themselves.

A derivation schema is not necessarily consistent (such as (7) below), as it may contain opposite literals in the form $x_i = \neg x_j$ for some i and j (this is in accordance with Definition 2).

Example 1. *Consider $\mathcal{T}_1$ such that $\mathcal{F}_1$, $\mathcal{S}_1$, $\mathcal{D}_1$ are as follows.*

$$\mathcal{F}_1 \begin{cases} p \\ q \end{cases} \qquad \mathcal{S}_1 \begin{cases} p \rightarrow s & (r_1) \\ q \rightarrow \neg s & (r_2) \\ p, s \rightarrow u & (r_3) \end{cases} \qquad \mathcal{D}_1 \begin{cases} \neg s \Rightarrow t & (r_4) \\ t, u \Rightarrow \neg v & (r_5) \\ p \Rightarrow q & (r_6) \end{cases}$$

Each of (1)–(7) below is a derivation schema from $\mathcal{T}_1$

$$\langle (p, \emptyset) \rangle \tag{1}$$

$$\langle (q, \emptyset), (\neg s, r_2) \rangle \tag{2}$$

$$\langle (p, \emptyset), (s, r_1), (u, r_3) \rangle \tag{3}$$

$$\langle (p, \emptyset), (s, r_1), (p, \emptyset), (u, r_3) \rangle \tag{4}$$

$$\langle (p, \emptyset), (q, \emptyset), (s, r_1), (u, r_3) \rangle \tag{5}$$

$$\langle (p, \emptyset), (q, r_6), (\neg s, r_2) \rangle \tag{6}$$

$$\langle (p, \emptyset), (q, \emptyset), (\neg s, r_2), (s, r_1), (u, r_3), (t, r_4), (\neg v, r_5) \rangle \tag{7}$$

A derivation schema may not be ($\subseteq$-)minimal. There are two reasons for that:

- repeating pairs (x_i, r_i) as in derivation (4) ($(p, \emptyset)$ is repeated twice),
- involving literals that do not serve towards inferring the conclusion x, as is illustrated by (5) (q is of no use there). The derivation schema fails thus to be *focussed*.

Definition 4 (Minimal/focussed derivation schema). *A derivation schema for x from $\mathcal{T}$ is* minimal *iff none of its proper subsequences is a derivation schema for x from $\mathcal{T}$. It is* focussed *iff it can be reduced to a minimal one by just deleting repeated pairs (x_i, r_i).*

Property 1. *Let $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$ be a theory. A derivation schema $d = \langle (x_1, r_1), \ldots, (x_n, r_n) \rangle$ from $\mathcal{T}$ is minimal iff d is focussed and the literals $x_1, \ldots, x_n$ are pairwise distinct.*

Notation. $\mathtt{CN}(\mathcal{T})$ denotes the set of all literals that have a derivation schema from $\mathcal{T}$. We call $\mathtt{CN}(\mathcal{T})$ the potential consequences drawn from $\mathcal{T}$ (for short, consequences) but they need not be definitive as they may be dismissed by opposite conclusions.

Property 2. *Let $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$ be a theory.*

- $\mathcal{F} \subseteq \texttt{CN}(\mathcal{T}) \subseteq \mathcal{F} \cup \{\texttt{Head}(r) \mid r \in \mathcal{S} \cup \mathcal{D}\} \subseteq \mathcal{L}$
- *If $\mathcal{T}$ is finite, then $\texttt{CN}(\mathcal{T})$ is finite*
- $\mathcal{F} = \emptyset$ *iff* $\texttt{CN}(\mathcal{T}) = \emptyset$
- *If d is a derivation schema from $\mathcal{T}$, $\texttt{Seq}(d) \subseteq \texttt{CN}(\mathcal{T})$*

Some rules may not be *activated* (i.e., their body has no derivation schema). Let us consider the following example.

Example 2. *Let $\mathcal{T}_2 = (\mathcal{F}_2, \mathcal{S}_2, \mathcal{D}_2)$ such that*

$$\mathcal{F}_2 \begin{cases} p \\ q \end{cases} \qquad \mathcal{S}_2 \begin{cases} p \to t & (r_1) \\ q \to t & (r_2) \\ s \to u & (r_3) \end{cases} \qquad \mathcal{D}_2 \begin{cases} p \Rightarrow q & (r_4) \\ u \Rightarrow v & (r_5) \end{cases}$$

There are rules here whose head is not a consequence of $\mathcal{T}_2$. $\texttt{CN}(\mathcal{T}_2) = \{p, q, t\} \subset \{p, q, t, u, v\} = \mathcal{F}_2 \cup \texttt{Head}(\mathcal{S}_2 \cup \mathcal{D}_2)$.

It is also easy to show that $\texttt{CN}$ is monotonic.

Property 3. *If $\mathcal{T} \sqsubseteq \mathcal{T}'$ then* $\texttt{CN}(\mathcal{T}) \subseteq \texttt{CN}(\mathcal{T}')$.

The backbone of an argumentation system is naturally the notion of *arguments*. They are built from a theory using the notion of derivation schema as follows.

Definition 5 (Argument). *Let $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$ be a theory. An* argument *defined from $\mathcal{T}$ is a pair (d, x) s.t.*

- $x \in \mathcal{L}$
- *d is a derivation schema for x from $\mathcal{T}$*
- $\texttt{Seq}(d)$ *is consistent*
- $\nexists \mathcal{T}' \sqsubset (\texttt{Facts}(d), \texttt{Strict}(d), \texttt{Def}(d))$ *s.t.* $x \in \texttt{CN}(\mathcal{T}')$

An argument (d, x) is strict *iff* $\texttt{Def}(d) = \emptyset$.

Notation. If $a = (d, x)$ is an argument then $\texttt{Conc}(a) = x$. For a set $\mathcal{E}$ of arguments, $\texttt{Concs}(\mathcal{E}) = \{x \mid (d, x) \in \mathcal{E}\}$. $\texttt{Arg}(\mathcal{T})$ is the set of all the arguments defined from $\mathcal{T}$. For a set $\mathcal{E}$ of arguments,

$$\texttt{Th}(\mathcal{E}) = (\bigcup_{(d,x)\in\mathcal{E}} \texttt{Facts}(d), \bigcup_{(d,x)\in\mathcal{E}} \texttt{Strict}(d), \bigcup_{(d,x)\in\mathcal{E}} \texttt{Def}(d)).$$

Theorem 1. *Let $\mathcal{T}$ be a theory. For all consistent sequence $d = \langle (x_1, r_1), \ldots, (x_n, r_n) \rangle$ from $\mathcal{T}$, the following two statements are equivalent:*

- *(d, x) is an argument (from $\mathcal{T}$)*
- *d is a focussed derivation schema from $\mathcal{T}$ s.t. $x = x_n$*

Definition 6 (Sub-argument). *An argument (d, x) is a* sub-argument *of (d', x') iff* $(\texttt{Facts}(d), \texttt{Strict}(d), \texttt{Def}(d)) \sqsubseteq (\texttt{Facts}(d'), \texttt{Strict}(d'), \texttt{Def}(d'))$.

Notation $\texttt{Sub}(a)$ denotes the set of all sub-arguments of a.

Example 1 (Cont). The argument $(\langle(q, \emptyset), (\neg s, r_2)\rangle, \neg s)$ has two sub-arguments: $(\langle(q, \emptyset)\rangle, q)$ and itself. By contrast, $(\langle(q, \emptyset)\rangle, q)$ is not a sub-argument of $(\langle(p, \emptyset), (q, r_6)\rangle, q)$.

Clearly, if (d, x) is a sub-argument of (d', x') then $\texttt{Seq}(d) \subseteq \texttt{Seq}(d')$, but the converse is not true as shown next.

Example 2 (Cont). Arguments $a = (\langle(p, \emptyset), (t, r_1)\rangle, t)$ and $b = (\langle(p, \emptyset), (q, r_4), (t, r_2)\rangle, t)$ are s.t. $\texttt{Seq}(a) = \{p, t\} \subseteq \{p, q, t\} = \texttt{Seq}(b)$ but a is not a sub-argument of b.

From the monotonicity of $\texttt{CN}$, it follows that the construction of arguments is a monotonic process.

Proposition 1. *If $\mathcal{T} \sqsubseteq \mathcal{T}'$ then $\texttt{Arg}(\mathcal{T}) \subseteq \texttt{Arg}(\mathcal{T}')$.*

An argumentation system is defined as follows:

Definition 7 (Argumentation system). *An* argumentation system *(AS for short) defined over a theory $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$ is a pair $\mathcal{H} = (\texttt{Arg}(\mathcal{T}), \mathcal{R})$ where $\mathcal{R} \subseteq \texttt{Arg}(\mathcal{T}) \times \texttt{Arg}(\mathcal{T})$ is called an attack relation.*

In what follows, arguments are evaluated using semantics proposed in [4]. Before recalling them, let us first introduce the two requirements on which they are based.

Definition 8 (Conflict-freeness – Defence). *Let $\mathcal{H} = (\mathcal{A}, \mathcal{R})$ be an AS, $\mathcal{E} \subseteq \mathcal{A}$ and $a \in \mathcal{A}$.*

- *$\mathcal{E}$ is* conflict-free *iff $\nexists a, b \in \mathcal{E}$ s.t. $a \mathcal{R} b$.*
- *$\mathcal{E}$* defends *a iff $\forall b \in \mathcal{A}$, if $b \mathcal{R} a$ then $\exists c \in \mathcal{E}$ s.t. $c \mathcal{R} b$.*

Definition 9 recalls the semantics of interest in the sequel.

Definition 9 (Acceptability semantics). *Let $\mathcal{H} = (\mathcal{A}, \mathcal{R})$ be an AS and $\mathcal{E} \subseteq \mathcal{A}$.*

- *$\mathcal{E}$ is a* naive *extensions iff it is a maximal (w.r.t. set $\subseteq$) conflict-free set.*
- *$\mathcal{E}$ is an* admissible *set iff it is conflict-free and defends all its elements.*
- *$\mathcal{E}$ is a* preferred *extension iff it is a maximal (w.r.t. set $\subseteq$) admissible set.*
- *$\mathcal{E}$ is a* stable *extension iff it is conflict-free and $\forall a \in \mathcal{A} \setminus \mathcal{E}$, $\exists b \in \mathcal{E}$ s.t. $b \mathcal{R} a$.*
- *$\mathcal{E}$ is a* grounded extension *iff it is a minimal (w.r.t. set $\subseteq$) set that is admissible and contains any argument it defends.*
- *$\mathcal{E}$ is an* ideal extension *iff it is the maximal (w.r.t. set $\subseteq$) admissible set which is part of any preferred extension.*

Notation. $\texttt{Ext}_x(\mathcal{H})$ denotes the set of all the extensions of a system $\mathcal{H}$ under semantics x where $x \in \{n, p, s\}$ and n (resp. p, s) stands for naive (resp. preferred, stable).

Plausible conclusions are those common to all extensions.

Definition 10 (Plausible conclusions). *If* $\mathcal{H} = (\texttt{Arg}(\mathcal{T}), \mathcal{R})$ *is an AS built over a theory* $\mathcal{T}$*, the set of* plausible conclusions *of* $\mathcal{H}$ *is*

$$\texttt{Output}(\mathcal{H}) = \bigcap_{\mathcal{E}_i \in \texttt{Ext}_x(\mathcal{H})} \texttt{Concs}(\mathcal{E}_i).$$

From the above definitions, namely that of an argument, it follows that the plausible conclusions of an argumentation system are a subset of the consequences that follow wrt CN from the theory over which the system is built.

Property 4. *Let* $\mathcal{H} = (\texttt{Arg}(\mathcal{T}), \mathcal{R})$ *be an AS built over* $\mathcal{T}$. $\texttt{Output}(\mathcal{H}) \subseteq \texttt{CN}(\mathcal{T})$.

3 Postulates for Argumentation Systems

We present rationality postulates that any rule-based argumentation system should satisfy. The first two were already proposed in [3] and the others are new. The first postulate ensures that the set of conclusions of arguments of each extension is consistent. This is compatible with the fact that each extension represents a coherent position.

Postulate 1 (Consistency). *Let* $\mathcal{H} = (\texttt{Arg}(\mathcal{T}), \mathcal{R})$ *be an AS built over a theory* $\mathcal{T}$*. For all* $\mathcal{E} \in \texttt{Ext}_x(\mathcal{H})$*,* $\texttt{Concs}(\mathcal{E})$ *is consistent. We say that* $\mathcal{H}$ *satisfies consistency.*

It was shown in [3] that if an argumentation system $\mathcal{H}$ satisfies consistency, then its set $\texttt{Output}(\mathcal{H})$ of plausible conclusions is consistent as well.

Property 5 ([3]). *If an AS* $\mathcal{H}$ *satisfies consistency, then* $\texttt{Output}(\mathcal{H})$ *is consistent.*

The second postulate ensures that the extensions of an argumentation system are closed under strict rules. The idea is that if there is an argument with conclusion x in an extension and there exists a strict rule $x \to y$, then y should also be supported by an argument in the same extension.

Postulate 2 (Closure under strict rules). *Let* $\mathcal{H} = (\texttt{Arg}(\mathcal{T}), \mathcal{R})$ *be an AS built over a theory* $\mathcal{T}$*. For all* $\mathcal{E} \in \texttt{Ext}_x(\mathcal{H})$*,* $\texttt{Concs}(\mathcal{E}) = \texttt{CN}((\texttt{Concs}(\mathcal{E}), \mathcal{S}, \emptyset))$*. We say that* $\mathcal{H}$ *is closed under strict rules.*

It is known that if an argumentation system $\mathcal{H}$ is closed under strict rules, then its set $\texttt{Output}(\mathcal{H})$ is necessarily closed under strict rules.

Property 6 ([3]). *Let* $\mathcal{H}$ *be an AS built over a theory* $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$*. If* $\mathcal{H}$ *is closed under strict rules, then* $\texttt{Output}(\mathcal{H}) = \texttt{CN}((\texttt{Output}(\mathcal{H}), \mathcal{S}, \emptyset))$.

It was also shown in [3] that a system that satisfies consistency and closure under strict rules satisfies *indirect* consistency.

Property 7 ([3]). *Let* $\mathcal{H}$ *be an AS built over a theory* $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$*. If* $\mathcal{H}$ *satisfies consistency and is closed under strict rules, then for all* $\mathcal{E} \in \texttt{Ext}_x(\mathcal{H})$*,* $\texttt{CN}((\texttt{Concs}(\mathcal{E}), \mathcal{S}, \emptyset))$ *is consistent.*

We propose three new postulates. The first says that if an argument belongs to an extension, then all its sub-arguments should be in the extension. It means that an argument cannot be accepted in an extension if one of its sub-parts is rejected.

Postulate 3 (Closure under sub-arguments). *Let $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ be an AS built over a theory $\mathcal{T}$. For all $\mathcal{E} \in \mathtt{Ext}_x(\mathcal{H})$, if $a \in \mathcal{E}$ then $\mathtt{Sub}(a) \subseteq \mathcal{E}$. We say that $\mathcal{H}$ is closed under sub-arguments.*

The following result characterizes the extensions of an argumentation system which is closed under sub-arguments.

Proposition 2. *If an AS $\mathcal{H}$ is closed under sub-arguments, then $\forall \mathcal{E} \in \mathtt{Ext}_x(\mathcal{H})$,*

- $\mathtt{Concs}(\mathcal{E}) = \mathtt{CN}(\mathtt{Th}(\mathcal{E}))$
- $\forall (d, x) \in \mathtt{Arg}(\mathtt{Th}(\mathcal{E}))$, $\mathtt{Seq}(d) \subseteq \mathtt{Concs}(\mathcal{E})$

Importantly, even when a system is closed under sub-arguments, the equality $\mathcal{E} = \mathtt{Arg}(\mathtt{Th}(\mathcal{E}))$ is not always true. This depends on the semantics as we will see later.

Proposition 3. *If an argumentation system $\mathcal{H}$ satisfies consistency and closure under sub-arguments, then $\forall \mathcal{E} \in \mathtt{Ext}_x(\mathcal{H})$, $\mathtt{CN}(\mathtt{Th}(\mathcal{E}))$ is consistent.*

Since facts and strict rules are the "hard" part in a theory, it is natural that any strict argument should be in all extensions. This principle is applied in default logic [9].

Postulate 4 (Strict precedence). *Let $\mathcal{H}$ be an AS built over a theory $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$. For all $\mathcal{E} \in \mathtt{Ext}_x(\mathcal{H})$, $\mathtt{Arg}((\mathcal{F}, \mathcal{S}, \emptyset)) \subseteq \mathcal{E}$. We say that $\mathcal{H}$ satisfies strict precedence.*

We show next that every argumentation system satisfying Postulate 4 infers all the conclusions that follow from the set of facts and the strict rules of a theory.

Proposition 4. *Let $\mathcal{H}$ be an AS built over a theory $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$. If $\mathcal{H}$ satisfies strict precedence, then $\mathcal{F} \subseteq \mathtt{CN}((\mathcal{F}, \mathcal{S}, \emptyset)) \subseteq \mathtt{Output}(\mathcal{H})$.*

Next is an important result for the rest of our study: it says that if an argumentation system over a theory $\mathcal{T}$ satisfies Postulates 2, 3, 4, then the set of literals deduced from $\mathtt{Th}(\mathcal{E})$, the theory of an extension $\mathcal{E}$, is exactly the one obtained from $\mathtt{Th}(\mathcal{E})$ extended by all facts and strict rules of $\mathcal{T}$ which are not in $\mathtt{Th}(\mathcal{E})$.

Proposition 5. *Let $\mathcal{H}$ be an argumentation system built over a theory $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$. If $\mathcal{H}$ satisfies postulates 2, 3, 4, then for all $\mathcal{E} \in \mathtt{Ext}_x(\mathcal{H})$,*

$$\mathtt{CN}(\mathtt{Th}(\mathcal{E})) = \mathtt{CN}((\mathcal{F}, \mathcal{S}, \bigcup_{(d,x) \in \mathcal{E}} \mathtt{Def}(d))).$$

The last postulate ensures a form of completeness of the extensions. It says that if the sequence of an argument is part of the conclusions of a given extension, then the argument (Definition 5 ensures consistency) should belong to the extension. Informally: If each step in the argument is good enough to be in the extension, then so is the argument itself.

Postulate 5 (Exhaustiveness). *Let $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ be an AS built over a theory $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$. For all $\mathcal{E} \in \mathtt{Ext}_x(\mathcal{H})$, for all $(d, x) \in \mathtt{Arg}(\mathcal{T})$, if $\mathtt{Seq}(d) \subseteq \mathtt{Concs}(\mathcal{E})$, then $(d, x) \in \mathcal{E}$.*

The extensions (under any semantics) of any argumentation system that satisfies exhaustiveness and closure under sub-arguments are closed in terms of arguments.

Proposition 6. *If an AS $\mathcal{H}$ is closed under sub-arguments and satisfies the exhaustiveness postulate, then $\forall \mathcal{E} \in \mathtt{Ext}_x(\mathcal{H})$, $\mathcal{E} = \mathtt{Arg}(\mathtt{Th}(\mathcal{E}))$.*

Under some semantics like naive and stable, Postulate 5 follows from consistency and closure under sub-arguments. This is mostly the case when the attack relation is *conflict-dependent*, that is, it captures the inconsistency of the theory over which the argumentation system is built.

Definition 11 (Conflict-dependency). *Let $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ be an argumentation system. The attack relation $\mathcal{R}$ is* conflict-dependent *iff for all $(d, x), (d', x') \in \mathtt{Arg}(\mathcal{T})$, if $(d, x)\ \mathcal{R}\ (d', x')$ then $\mathtt{Seq}(d) \cup \mathtt{Seq}(d')$ is inconsistent.*

Proposition 7. *Let $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ be an argumentation system built over a theory $\mathcal{T}$ s.t. $\mathcal{R}$ is conflict-dependent. If $\mathcal{H}$ satisfies consistency and closure under sub-arguments, then $\mathcal{H}$ satisfies exhaustiveness under naive and stable semantics.*

Finally, it is worth noticing that conflict-dependent relations do not admit self-attacking arguments.

Proposition 8. *Let $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ be an argumentation system. If $\mathcal{R}$ is conflict-dependent, $\forall a \in \mathtt{Arg}(\mathcal{T})\ (a, a) \notin \mathcal{R}$.*

4 Outcomes of Argumentation Systems

This section analyzes the outputs of rule-based argumentation systems under the semantics recalled in Def. 9. In the sequel, we consider *only* systems that satisfy the postulates introduced in Section 3. As in [3,5,9], we assume that the "hard" part of a theory is consistent. Formally:

Assumption 2. *For all theory $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$, $\mathtt{CN}((\mathcal{F}, \mathcal{S}, \emptyset))$ is consistent.*

Let us first introduce a key concept: that of an *option*.

Definition 12 (Option). *Let $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$ be a theory. An* option *of $\mathcal{T}$ is a sub-theory $\mathcal{T}' = (\mathcal{F}', \mathcal{S}', \mathcal{D}')$ of $\mathcal{T}$ such that:*

- *$\mathcal{F}' = \mathcal{F}$ and $\mathcal{S}' = \mathcal{S}$ (hence $\mathcal{D}' \subseteq \mathcal{D}$)*
- *$\mathtt{CN}(\mathcal{T}')$ is consistent*
- *$\forall r \in \mathcal{D} \setminus \mathcal{D}'$, $\mathtt{CN}((\mathcal{F}, \mathcal{S}, \mathcal{D}' \cup \{r\}))$ is inconsistent.*

Let $\mathtt{Opt}(\mathcal{T})$ denote the set of all options of $\mathcal{T}$.

Example 3. *Consider $\mathcal{T}_3$ such that $\mathcal{F}_3$, $\mathcal{S}_3$, $\mathcal{D}_3$ are as follows.*

$$\mathcal{F}_3 \begin{cases} p \\ q \\ \neg s \end{cases} \qquad \mathcal{S}_3 \left\{ t, u, v \to s \quad (r_1) \right. \qquad \mathcal{D}_3 \begin{cases} p \Rightarrow t & (r_2) \\ q \Rightarrow u & (r_3) \\ u \Rightarrow v & (r_4) \end{cases}$$

The theory $\mathcal{T}_3$ has three options:

- $\mathcal{O}_1 = (\mathcal{F}_3, \mathcal{S}_3, \{p \Rightarrow t, q \Rightarrow u\})$
- $\mathcal{O}_2 = (\mathcal{F}_3, \mathcal{S}_3, \{p \Rightarrow t, u \Rightarrow v\})$
- $\mathcal{O}_3 = (\mathcal{F}_3, \mathcal{S}_3, \{q \Rightarrow u, u \Rightarrow v\})$

When a theory is consistent, it has a unique option: itself. This is the case in Example 2: $\texttt{Opt}(\mathcal{T}_2) = \{\mathcal{T}_2\}$.

Property 8. *Let $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$ be a theory.*

- $\texttt{Opt}(\mathcal{T}) = \{\mathcal{T}\}$ *iff* $\texttt{CN}(\mathcal{T})$ *is consistent.*
- *If* $\texttt{CN}((\mathcal{F}, \mathcal{S}, \emptyset))$ *is inconsistent, then* $\texttt{Opt}(\mathcal{T}) = \emptyset$.
- *For all $r \in \mathcal{D}$, if $\texttt{CN}((\mathcal{F}, \mathcal{S}, \{r\}))$ is consistent, then there exists an option $\mathcal{O}$ s.t. $(\mathcal{F}, \mathcal{S}, \{r\}) \sqsubseteq \mathcal{O}$.*

Note that the set of consequences of an option is not necessarily maximal for set inclusion as shown by Example 3.

Example 3 (Cont). We have $\texttt{CN}(\mathcal{O}_1) = \{p, q, \neg s, t, u\}$, $\texttt{CN}(\mathcal{O}_2) = \{p, q, \neg s, t\}$, and $\texttt{CN}(\mathcal{O}_3) = \{p, q, \neg s, u, v\}$. Thus, $\texttt{CN}(\mathcal{O}_2) \subseteq \texttt{CN}(\mathcal{O}_1)$.

Notation For a set $\mathcal{B}$ of theories, we denote its maximum as $\texttt{Max}(\mathcal{B}) = \{\mathcal{T} \in \mathcal{B} \mid \nexists \mathcal{T}' \in \mathcal{B}$ s.t. $\texttt{CN}(\mathcal{T}) \subset \texttt{CN}(\mathcal{T}')\}$. In Example 3, $\texttt{Max}(\texttt{Opt}(\mathcal{T}_3)) = \{\mathcal{O}_1, \mathcal{O}_3\}$.

The defeasible rules of a theory do not necessarily belong to an option of the theory as shown by the following example.

Example 4. *The theory $\mathcal{T}_4$ s.t. $\mathcal{F}_4 = \{p, q\}$, $\mathcal{S}_4 = \{p \to s\}$ and $\mathcal{D}_4 = \{q \Rightarrow \neg s\}$ has a single option: $\mathcal{O} = (\mathcal{F}_4, \mathcal{S}_4, \emptyset)$.*

4.1 Naive Semantics

We start by characterizing the naive extensions of any argumentation system satisfying the above rationality postulates. We show that each naive extension returns a maximal option of the theory over which the system is built.

Theorem 2. *Let $\mathcal{H} = (\texttt{Arg}(\mathcal{T}), \mathcal{R})$ be an AS built over a theory $\mathcal{T}$ s.t. $\mathcal{R}$ is conflict-dependent and $\mathcal{H}$ satisfies the postulates 1, 2, 3, and 4. For all $\mathcal{E} \in \texttt{Ext}_n(\mathcal{H})$, there exists a unique $\mathcal{O} \in \texttt{Max}(\texttt{Opt}(\mathcal{T}))$ such that $\texttt{Th}(\mathcal{E}) \sqsubseteq \mathcal{O}$ and $\texttt{Concs}(\mathcal{E}) = \texttt{CN}(\mathcal{O})$.*

Note that the theory of a naive extension may be a proper subset of the corresponding maximal option. This is mainly due to the fact that an option may contain non-activated rules while arguments are minimal and thus focussed.

Example 2 (Cont). Since theory $\mathcal{T}_2$ is consistent, then it has a single (maximal) option which is the theory itself. Any AS built over $\mathcal{T}_2$ and which obeys the postulates and whose attack relation is conflict-dependent will have a single naive extension $\mathcal{E}$ with $\mathtt{Th}(\mathcal{E}) = (\mathcal{F}, \{r_1, r_2\}, \{r_5\}) \sqsubset \mathcal{T}_2$. Rules r_3 and r_5 are not used in arguments.

Notation For $\mathcal{E}$ naive extension of $\mathcal{H}$ s.t. $\mathcal{O}$ in $\mathtt{Max}(\mathtt{Opt}(\mathcal{T}))$ satisfies $\mathtt{Th}(\mathcal{E}) \sqsubseteq \mathcal{O}$ and $\mathtt{Concs}(\mathcal{E}) = \mathtt{CN}(\mathcal{O})$, let $\mathtt{Option}(\mathcal{E}) \stackrel{\text{def}}{=} \mathcal{O}$.

We prove that no two naive extensions return the same option. Moreover, naive extensions are closed in terms of arguments.

Theorem 3. *Let $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ be an AS built over a theory $\mathcal{T}$ s.t. $\mathcal{R}$ is conflict-dependent and $\mathcal{H}$ satisfies the postulates 1, 2, 3, and 4.*

- *For all $\mathcal{E}, \mathcal{E}' \in \mathtt{Ext}_n(\mathcal{H})$, if $\mathtt{Option}(\mathcal{E}) = \mathtt{Option}(\mathcal{E}')$, then $\mathcal{E} = \mathcal{E}'$*
- *For all $\mathcal{E} \in \mathtt{Ext}_n(\mathcal{H})$, $\mathcal{E} = \mathtt{Arg}(\mathtt{Option}(\mathcal{E}))$*

We have shown that each naive extension captures exactly one maximal option and it supports all, and only, the consequences of that option. Theorem 4 states that every option has a corresponding naive extension. So, there is a bijection from the set of naive extensions to the set of maximal options.

Theorem 4. *Let $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ be an AS built over a theory $\mathcal{T}$ s.t. $\mathcal{R}$ is conflict-dependent and $\mathcal{H}$ satisfies the postulates 1, 2, 3, and 4.*

- *For all $\mathcal{O} \in \mathtt{Max}(\mathtt{Opt}(\mathcal{T}))$, $\mathtt{Arg}(\mathcal{O}) \in \mathtt{Ext}_n(\mathcal{H})$.*
- *For all $\mathcal{O} \in \mathtt{Max}(\mathtt{Opt}(\mathcal{T}))$, $\mathcal{O} = \mathtt{Option}(\mathtt{Arg}(\mathcal{O}))$*
- *For all $\mathcal{O}, \mathcal{O}' \in \mathtt{Max}(\mathtt{Opt}(\mathcal{T}))$, if $\mathtt{Arg}(\mathcal{O}) = \mathtt{Arg}(\mathcal{O}')$ then $\mathcal{O} = \mathcal{O}'$.*

Example 3 (Cont). The theory $\mathcal{T}_3$ has three options, of which only two are maximal: $\mathtt{Max}(\mathtt{Opt}(\mathcal{T})) = \{\mathcal{O}_1, \mathcal{O}_3\}$. For all argumentation system $\mathcal{H}$ built over $\mathcal{T}_3$, if the attack relation of $\mathcal{H}$ is to be conflict-dependent and the postulates satisfied, then $\mathtt{Ext}_n(\mathcal{H}) = \{\mathtt{Arg}(\mathcal{O}_1), \mathtt{Arg}(\mathcal{O}_3)\}$.

It is thus possible to delimit the number of naive extensions of any argumentation system that satisfies the four postulates.

Corollary 1. *Let $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ be an AS built over a theory $\mathcal{T}$ s.t. $\mathcal{R}$ is conflict-dependent and $\mathcal{H}$ satisfies the postulates 1, 2, 3, and 4. The equality $|\mathtt{Ext}_n(\mathcal{H})| = |\mathtt{Max}(\mathtt{Opt}(\mathcal{T}))|$ holds.*

What about the plausible conclusions that are drawn from a theory using an argumentation system that satisfies the postulates? From the previous results, it is easy to show that they are the literals that follow from all the maximal options.

Theorem 5. *Let $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ be an AS built over a theory $\mathcal{T}$ s.t. $\mathcal{R}$ is conflict-dependent and $\mathcal{H}$ satisfies the postulates 1, 2, 3, and 4.*

$$\mathtt{Output}(\mathcal{H}) = \bigcap_{\mathcal{O}_i \in \mathtt{Max}(\mathtt{Opt}(\mathcal{T}))} \mathtt{CN}(\mathcal{O}_i)$$

Example 3 (Cont). Any argumentation system $\mathcal{H}$ that can be built over the theory $\mathcal{T}_3$ and has a conflict-dependent attack relation and satisfies the postulates 1, 2, 3, 4 will have as output the set $\mathtt{Output}(\mathcal{H}) = \mathtt{CN}(\mathcal{O}_1) \cap \mathtt{CN}(\mathcal{O}_2) = \{p, q, \neg s, u\}$.

4.2 Stable Semantics

We now analyze the outcomes of rule-based argumentation systems under stable semantics, again considering only systems that satisfy the rationality postulates. We show that such systems have stable extensions if the set of facts is not empty.

Theorem 6. *Let* $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$ *be a theory.Whenever* $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ *is an AS satisfying postulate 4,* $|\mathtt{Ext}_s(\mathcal{H})| = 0$ *iff* $\mathcal{F} = \emptyset$*.*

As for naive extensions, stable extensions of any argumentation system that satisfies the postulates return maximal options of the theory at hand.

Theorem 7. *Let* $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ *be an AS defined over a theory* $\mathcal{T}$ *s.t.* $\mathcal{R}$ *is conflict-dependent and* $\mathcal{H}$ *satisfies the postulates 1, 2, 3, 4. For all* $\mathcal{E} \in \mathtt{Ext}_s(\mathcal{H})$*,* $\exists!\mathcal{O} \in \mathtt{Max}(\mathtt{Opt}(\mathcal{T}))$ *s.t.*

- $\mathtt{Th}(\mathcal{E}) \sqsubseteq \mathcal{O}$ *and* $\mathtt{Concs}(\mathcal{E}) = \mathtt{CN}(\mathcal{O})$.
- $\mathcal{E} = \mathtt{Arg}(\mathcal{O})$.

Two stable extensions capture distinct options.

Theorem 8. *Let* $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ *be an AS defined over a theory* $\mathcal{T}$ *s.t.* $\mathcal{R}$ *is conflict-dependent and* $\mathcal{H}$ *satisfies the postulates 1, 2, 3, 4.*
For all $\mathcal{E}, \mathcal{E}' \in \mathtt{Ext}_s(\mathcal{H})$*, if* $\mathtt{Option}(\mathcal{E}) = \mathtt{Option}(\mathcal{E}')$ *then* $\mathcal{E} = \mathcal{E}'$*.*

Corollary 2. *Let* $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ *be an AS defined over a theory* $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$ *s.t.* $\mathcal{F} \neq \emptyset$ *and* $\mathcal{R}$ *is conflict-dependent and* $\mathcal{H}$ *satisfies Postulates 1,2,3,4. It holds that* $1 \leq |\mathtt{Ext}_s(\mathcal{H})| \leq |\mathtt{Max}(\mathtt{Opt}(\mathcal{T}))|$*.*

Theorem 7 does not guarantee that each maximal option of a theory $\mathcal{T}$ has a corresponding stable extension. The equality $|\mathtt{Ext}_s(\mathcal{H})| = |\mathtt{Max}(\mathtt{Opt}(\mathcal{T}))|$ depends on the attack relation. Let $\Re_s$ be the set of *all* attack relations that are conflict-dependent and that ensure Postulates 1, 2, 3, 4 under stable semantics. This set contains two *disjoints* subsets of attack relations, i.e. $\Re_s = \Re_{s_1} \cup \Re_{s_2}$:

- $\Re_{s_1}$: the relations s.t. $|\mathtt{Ext}_s(\mathcal{H})| < |\mathtt{Max}(\mathtt{Opt}(\mathcal{T}))|$
- $\Re_{s_2}$: the relations s.t. $|\mathtt{Ext}_s(\mathcal{H})| = |\mathtt{Max}(\mathtt{Opt}(\mathcal{T}))|$

Systems that use relations in $\Re_{s_1}$ choose a proper subset of the maximal options of $\mathcal{T}$ and make inferences from them. Their output sets are as follows:

Theorem 9. *Let* $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ *be an argumentation system built over a theory* $\mathcal{T}$ *s.t.* $\mathcal{R} \in \Re_{s_1}$. $\mathtt{Output}(\mathcal{H}) = \bigcap_{\mathcal{O}_i \in \mathcal{S}} \mathtt{CN}(\mathcal{O}_i)$ *with* $\mathcal{S} = \{\mathcal{O}_i \in \mathtt{Max}(\mathtt{Opt}(\mathcal{T})) \mid \mathtt{Arg}(\mathcal{O}_i) \in \mathtt{Ext}_s(\mathcal{H})\}$.

These attack relations introduce a "critical discrimination" between the maximal options of a theory. Hence, great care must be exercised when designing rule-based argumentation systems based on stable semantics: The principles governing the interaction between $\Rightarrow$ and $\mathcal{R}$ must be both rigorously and meticulously specified so as to avoid trouble of which the following example is an easy case.

Example 5. *Let $\mathcal{T}_5$ be s.t. $\mathcal{F}_5 = \{p, q\}$ and $\mathcal{S}_5 = \emptyset$ and $\mathcal{D}_5 = \{p \Rightarrow s, q \Rightarrow \neg s\}$. $\mathcal{T}_5$ has two maximal options: $\mathcal{O}_1 = (\mathcal{F}_5, \mathcal{S}_5, \{p \Rightarrow s\})$ and $\mathcal{O}_2 = (\mathcal{F}_5, \mathcal{S}_5, \{q \Rightarrow \neg s\})$. For any system $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}_5), \mathcal{R})$ s.t. $\mathcal{R} \in \Re_{s_1}$, either i) $\mathtt{Ext}_s(\mathcal{H}) = \{\mathtt{Arg}(\mathcal{O}_1)\}$ or ii) $\mathtt{Ext}_s(\mathcal{H}) = \{\mathtt{Arg}(\mathcal{O}_2)\}$. In case (i), $s \in \mathtt{Output}(\mathcal{H})$ and $\neg s \notin \mathtt{Output}(\mathcal{H})$. In case (ii), $\neg s$ is the plausible conclusion. Either choice would be arbitrary.*

Attack relations of category $\Re_{s2}$ induce a bijection between the stable extensions of an argumentation system and the maximal options of the theory over which it is built.

Theorem 10. *Let $\mathcal{T}$ = $(\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ be an argumentation system over a theory $\mathcal{T}$ s.t. $\mathcal{R} \in \Re_{s2}$. For all $\mathcal{O} \in \mathtt{Max}(\mathtt{Opt}(\mathcal{T}))$, $\mathtt{Arg}(\mathcal{O}) \in \mathtt{Ext}_s(\mathcal{H})$.*

Argumentation systems with an attack relation from $\Re_{s2}$ are *coherent*, meaning that the preferred extensions exhaust all and only the stable ones.

Theorem 11. *Let $\mathcal{T}$ = $(\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ be an argumentation system over a theory $\mathcal{T}$ s.t. $\mathcal{R} \in \Re_{s2}$. $\mathtt{Ext}_s(\mathcal{H}) = \mathtt{Ext}_p(\mathcal{H}) = \mathtt{Ext}_n(\mathcal{H})$.*

Attack relations in category $\Re_{s2}$ conform exactly to the result obtained under naive semantics: Plausible conclusions for them are already characterized in Theorem 5.

To sum up, attack relations satisfying the postulates can be split into two categories: $\Re_{s_1}$ and $\Re_{s_2}$. Relations from $\Re_{s_2}$ do not offer added value as they make the stable semantics case to collapse to the naive semantics case. For stable semantics to substantiate (as compared with naive semantics) a rule-based argumentation system, attack relations from category $\Re_{s_1}$ must be favored. However, pitfalls threaten as options are discarded, and a lot of care must be exercised when designing such a system.

4.3 Preferred Semantics

Preferred semantics was initially proposed to overcome the limitation of stable semantics which does not guarantee the existence of extensions. Indeed, any argumentation system has at least one preferred extension which may be empty. We show that in case of rule-based systems the empty set cannot be an extension.

Proposition 9. *Let $\mathcal{H}$ be an AS built over a theory $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$ s.t. $\mathcal{H}$ satisfies strict precedence. $\mathtt{Ext}_p(\mathcal{H}) = \{\emptyset\}$ iff $\mathcal{F} = \emptyset$.*

Unlike the cases of naive and stable extensions, a preferred extension may capture only a sub-part of the consequences drawn from a maximal option.

Theorem 12. *Let $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ be an AS built over a theory $\mathcal{T}$ s.t. $\mathcal{R}$ is conflict-dependent and $\mathcal{H}$ satisfies the postulates 1 and 3. For all $\mathcal{E} \in \mathtt{Ext}_p(\mathcal{H})$, $\exists \mathcal{O} \in \mathtt{Max}(\mathtt{Opt}(\mathcal{T}))$ s.t. $\mathtt{Th}(\mathcal{E}) \sqsubseteq \mathcal{O}$ and $\mathtt{Concs}(\mathcal{E}) \subseteq \mathtt{CN}(\mathcal{O})$.*

Two preferred extensions refer to different options.

Theorem 13. *Let $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ be an argumentation system s.t. $\mathcal{R}$ is conflict-dependent and $\mathcal{H}$ satisfies the postulates 1, 2, 3, and 4. Let $\mathcal{E}, \mathcal{E}' \in \mathtt{Ext}_p(\mathcal{H})$ and $\mathcal{O} \in \mathtt{Max}(\mathtt{Opt}(\mathcal{T}))$. If $\mathtt{Th}(\mathcal{E}) \sqsubseteq \mathcal{O}$ and $\mathtt{Th}(\mathcal{E}') \sqsubseteq \mathcal{O}$, then $\mathcal{E} = \mathcal{E}'$.*

From the previous result, it follows that the number of preferred extensions does not exceed the number of maximal options of the theory over which the system is built.

Theorem 14. *Let* $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ *be a system built over a theory* $\mathcal{T}$ *s.t.* $\mathcal{R}$ *is conflict-dependent and* $\mathcal{H}$ *satisfies Postulates 1, 2, 3, and 4.* $|\mathtt{Ext}_p(\mathcal{H})| \leq |\mathtt{Max}(\mathtt{Opt}(\mathcal{T}))|$.

Regarding the outputs of a rule-based argumentation system under preferred semantics, there are two cases: i) Attack relations of category $\mathcal{R}_{s2}$ lead to coherent systems whose plausible conclusions are characterized by Theorem 5. Thus, naive, stable and preferred semantics coincide. ii) Attack relations of category $\mathcal{R}_{s1}$ lead to pick up some maximal options and to reason about them. The plausible conclusions are given by Theorem 9. Thus the situation about preferred semantics is similar with that for stable semantics: For preferred semantics to offer added value over naive semantics, the attack relation chosen must discard some maximal options but it takes a lot of care to specify such an attack relation in full generality.

4.4 Grounded Semantics – Ideal Semantics

This section analyses the outcomes of rule-based systems under grounded and ideal semantics. We show that the ideal extension is exactly the set of arguments built from the *free* part of a theory. The free part of a theory $\mathcal{T} = (\mathcal{F}, \mathcal{S}, \mathcal{D})$, denoted by $\mathtt{Free}(\mathcal{T})$, is a sub-theory $(\mathcal{F}, \mathcal{S}, \mathcal{D}')$ where $\mathcal{D}' = \cap \mathcal{D}_i$ where $(\mathcal{F}, \mathcal{S}, \mathcal{D}_i) \in \mathtt{Opt}(\mathcal{T})$. In other words, $\mathcal{D}'$ contains all the defeasible rules that are not involved in any conflict.

Proposition 10. *Let* $\mathcal{T}$ *be a theory.* $\mathtt{CN}(\mathtt{Free}(\mathcal{T}))$ *is consistent.*

We show that when the attack relation satisfies a very natural requirement, then $\mathtt{Arg}(\mathtt{Free}(\mathcal{T}))$ is admissible (i.e., it is conflict-free and defends all its elements).

Definition 13. *Let* $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ *be an AS over a theory* $\mathcal{T}$*. An attack relation* $\mathcal{R}$ privileges strict arguments *iff for all* $a = (d, x), b = (d', x') \in \mathtt{Arg}(\mathcal{T})$*, if* a *is strict and* $\mathtt{Seq}(d) \cup \mathtt{Seq}(d')$ *is inconsistent, then* $a\mathcal{R}b$*.*

As far as we know, all the attack relations in existing rule-based argumentation systems privilege strict arguments.

Theorem 15. *Let* $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ *be a system built over a theory* $\mathcal{T}$ *s.t.* $\mathcal{R}$ *is conflict-dependent and privileges strict arguments.* $\mathtt{Arg}(\mathtt{Free}(\mathcal{T}))$ *is admissible.*

The set $\mathtt{Arg}(\mathtt{Free}(\mathcal{T}))$ is part of every preferred extension.

Theorem 16. *Let* $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ *be an AS over a theory* $\mathcal{T}$ *s.t.* $\mathcal{R}$ *is conflict-dependent and privileges strict arguments, and* $\mathcal{H}$ *satisfies Postulates 1, 3, 4.*
$\mathtt{Arg}(\mathtt{Free}(\mathcal{T})) \subseteq \bigcap_{\mathcal{E}_i \in \mathtt{Ext}_p(\mathcal{H})} \mathcal{E}_i$.

We show next that in case of attack relations of category $\Re_{s2}$, $\mathtt{Arg}(\mathtt{Free}(\mathcal{T}))$ is equal to the intersection of all preferred extensions. Recall that in this case, preferred extensions coincide with stable extensions and with naive ones.

Theorem 17. *Let $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ be an AS over a theory $\mathcal{T}$. If $\mathcal{R} \in \Re_{s2}$ then* $\mathtt{Arg}(\mathtt{Free}(\mathcal{T})) = \bigcap_{\mathcal{E}_i \in \mathtt{Ext}_x(\mathcal{H})} \mathcal{E}_i$.

From the previous result, it follows that when the attack relation is of category $\Re_{s2}$ and privileges strict arguments, then $\mathtt{Arg}(\mathtt{Free}(\mathcal{T}))$ is the ideal extension.

Theorem 18. *Let $\mathcal{H} = (\mathtt{Arg}(\mathcal{T}), \mathcal{R})$ be an AS over a theory $\mathcal{T}$. If $\mathcal{R} \in \Re_{s2}$ and privileges strict arguments, then*

- $\mathtt{Arg}(\mathtt{Free}(\mathcal{T}))$ *is the ideal extension of $\mathcal{H}$.*
- *The grounded extension of $\mathcal{H}$ is a subset of* $\mathtt{Arg}(\mathtt{Free}(\mathcal{T}))$.

The above result shows that ideal and grounded semantics allow the inference of literals only from the free part of a theory. Note also that grounded extension is more cautious than ideal one and may miss intuitive (free) conclusions.

5 Conclusion

The paper provides the first investigation on the outputs of rule-based argumentation systems. The study is general in the sense that it keeps the attack relation unspecified. Thus, the system can be instantiated with any of the attack relations that are used in existing systems. The results show that under naive semantics, the systems return the literals that follow from all the options of the theory at hand. Stable and preferred semantics either do not provide an added value wrt naive semantics or the attack relation of a system should be formalized in a very rigorous way in order to avoid arbitrary results. Ideal semantics returns the free part of a theory whereas the grounded semantics returns a sub-part of the free part meaning that it may miss interesting conclusions.

References

1. Amgoud, L., Besnard, P.: Logical limits of abstract argumentation frameworks. Journal of Applied Non-Classical Logics (2013)
2. Amgoud, L., Cayrol, C.: Inferring from inconsistency in preference-based argumentation frameworks. Inter. J. of Automated Reasoning 29(2), 125–169 (2002)
3. Caminada, M., Amgoud, L.: On the evaluation of argumentation formalisms. Artificial Intelligence J. 171(5-6), 286–310 (2007)
4. Dung, P.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. AI. J. 77(2), 321–357 (1995)
5. García, A., Simari, G.: Defeasible logic programming: an argumentative approach. Theory and Practice of Logic Programming 4(1-2), 95–138 (2004)
6. Gorogiannis, N., Hunter, A.: Instantiating abstract argumentation with classical logic arguments: Postulates and properties. Artificial Intelligence J. 175(9-10), 1479–1497 (2011)
7. Governatori, G., Maher, M., Antoniou, G., Billington, D.: Argumentation semantics for defeasible logic. J. of Logic and Computation 14(5), 675–702 (2004)
8. Marek, V., Nerode, A., Remmel, J.: A theory of nonmonotonic rule systems I. Annals of Mathematics and Artificial Intelligence 1, 241–273 (1990)
9. Reiter, R.: A logic for default reasoning. Artificial Intelligence J. 13(1-2), 81–132 (1980)